AFRL-OSR-VA-TR-2013-0500

DISTRIBUTED CONTROL AND INFORMATION FUSION OVER COMMUNICATION NETWORKS

RAHUL JAIN

UNIVERSITY OF SOUTHERN CALIFORNIA

09/17/2013
Final Report

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSL
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
13-09-2013	Final Techincal Report	06/15/2010 - 06/14/2013

4. TITLE AND SUBTITLE

Distributed Control and Information Fusion Over Communication Networks

5a. CONTRACT NUMBER

5b. GRANT NUMBER

FA9550-10-1-0307

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Rahul Jain

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Southern California
3720 S. Flower St. CUB 305
Los Angeles, CA 90089-0701

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Office of Scientific Research
875 N. Randolph St. Room 3122
Arlington VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution is Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project focused on distributed control and information fusion/learning over communication systems. The first set of problems considered were related to distributed adaptive control. Namely, a first such decentralized learning algorithm for multi-armed bandit models was developed that achieved poly-logarithmic regret. Later another algorithm was given that achieve log regret. The second set of problems was related to decentralized control design for LQG systems with asymmetric one-step delayed information sharing between two controllers. The optimal decentralized controller was computed. A case with partial output feedback was also solved. The third problem solved was development of empirical value iteration, a new method for approximate dynamic programming. Convergence of such an algorithm to the optimal policy was established by introducing new notions of probabilistic fixed points of random operators. Numerically, this new class of methods is found to perform better than reinforcement learning methods.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Rahul Jain
UU	UU	UU	SAR	2	**19b. TELEPHONE NUMBER** *(include area code)* (213) 740-2246

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Final Performance Report, September 13, 2013

DISTRIBUTED CONTROL AND INFORMATION FUSION OVER COMMUNICATION NETWORKS
Contract Number: FA9550-10-1-0307

To: technicalreports@afosr.af.mil
Subject: Final Performance Report for Dr. Tristan Nguyen

Contract/Grant Title: Distributed Control and Information Fusion Over Communication Networks
Contract/Grant #: FA9550-10-1-0307
Reporting Period: 15 June 2010 - 14 June 2013

Project Accomplishments:

This project focused on distributed control and information fusion/decentralized learning over communication systems. The first set of problems we considered related to distributed adaptive control. These are formulated as multi-armed bandit problems. A decentralized learning algorithm, called dUCB4, for multi-player multi-armed bandit models was proposed that had poly-log regret. This is the first decentralized learning algorithm with sublinear regret. Later, another algorithm, the phased exploration and exploitation (PEE) algorithm was proposed that builds on dUCB4, and has improved performance with log regret. Two conference papers have been published and a journal paper on this is under review at the IEEE Trans. on Information Theory. Another is under preparation for submission to IEEE Trans. on Automatic Control. This has led us to addressing the problem when there are multiple controller, each acting autonomously. This is formulated as a multi-armed bandit game. Some progress has been made on this problem as well.

The second set of problems we considered relate to decentralized control in multi-user communication networks. A model was considered wherein there are multiple controllers who exchange their information with other controllers but with some delay. This is a classical decentralized control problem that has remained open for over 40 years. An asymmetric delayed sharing pattern was considered wherein it is known that they have a `partially nested information structure' and an linear optimal controller exists. The effort was on computing this linear optimal controller. Progress on this class of problems has been stalled for more than 30 years, until recently. Two important information sharing pattern problems were solved. A submitted journal paper has resulted along with two conference papers.

The third set of problems we considered was to develop a new approximate dynamic programming method, which we call the `empirical dynamic programming (EDP)' algorithm. Herein, any expectation is replaced by a sample average approximation. The open question has been whether such an algorithm will converge, and if so, whether it will converge to the optimal policy. We have been able to prove convergence to the optimal policy in a probabilistic sense. This required new conceptual developments on probabilistic fixed points of random operators. Numerical results show that EDP performs better than reinforcement learning and other stochastic approximations methods. Two journal paper are to be submitted to Mathematics of Operations Research and Automatica respectively soon. In addition, two conference papers will be submitted later.

The project has involved and supported two PhD students (Dileep Kalathil and Naumaan Nayyar), partially supported two postdocs (Arman Khouzani and Will Haskell) and the PI. Dileep Kalathil's PhD dissertation is expected to be submitted in May 2013, and will largely be a compilation of papers [1], [3], [4], [5] - work done under this project. Naumaan's dissertation is expected to be completed by May 2014,

and will include results in [2]. In all, the project has resulted in 5 journal publications/submissions, and 4 conference papers/submissions so far.

Archival publications (published/submitted) during reporting period:

1. D. Kalathil, N. Nayyar and R. Jain, "Decentralized learning for multi-player multi-armed bandits", submitted to, *IEEE Trans. on Information Theory*, April 2012. (third revision submitted: August 2013)

2. N. Nayyar, D. Kalathil and R. Jain, "Optimal decentralized control for asymmetric delayed information sharing", submitted to *IEEE Trans. on Automatic Control*, July 2013. (preprint available)

3. W. Haskell, D. Kalathil and R. Jain, "Empirical value iteration: an approximate dynamic programming method", in preparation for, *Mathematics of Operations Research,* 2013. (preprint available)

4. D. Kalathil, W. Haskell and R. Jain, "Empirical dynamic programming with fixed samples", in preparation for, *Automatica*, 2013. (preprint available)

5. D. Kalathil, N. Nayyar and R. Jain, "Phased Exploration and Exploitation Algorithm for multi-player multi-armed bandits", in preparation for, *IEEE Trans. on Automatic Control*, 2013.

6. D. Kalathil, N. Nayyar and R. Jain, "Decentralized learning for multi-player multi-armed bandits", *Proc. of the IEEE Conf. on Decision and Control (CDC)*, Dec 2012.

7. D. Kalathil, N. Nayyar and R. Jain, "Multi-player multi-armed bandits: Decentralized learning with IID rewards", *Proc. of The Allerton Conference*, Oct. 2012.

8. N. Nayyar, D. Kalathil and R. Jain, "Optimal decentralized control in unidirectional one-step delayed sharing pattern", to appear, *Proc. of The Allerton Conference*, Oct. 2013.

9. N. Nayyar, D. Kalathil and R. Jain, "Optimal decentralized control in unidirectional one-step delayed sharing with partial output feedback", for submission to, *The American Control Conference*, Sep. 2013.

Changes in research objectives, if any: None

Change in AFOSR program manager, if any: Initially, Dr. Doug Cochran. Later, Dr. Tristan Nguyen.

Extensions granted or milestones slipped, if any: None

Include any new discoveries, inventions, or patent disclosures during this reporting period (if none, report none): None